

WO 02/074943

PCT/FR02/00951

SEQUENCE LISTING

<110> Centre National de la Recherche Scientifique

<120> NUCLEIC ACID MOLECULES ENCODING A DEXTRANSUCRASE CATALYZING THE
SYNTHESIS OF DEXTRAN CARRYING ALPHA-1,2-OSIDIC TYPE LINKAGES

<130> B4787A (INPI) CNRS/INSA TOULOUSE

<140> xxxxxxxx

<141> 2001-12-19

<150> 0103631

<151> 2001-03-16 .

<160> 17

<170> PatentIn Ver. 2.1

<210> 1

<211> 855

<212> PRT

<213> Amino acids (catalytic domain n° 2)

<400> 1

Asp	Met	Ser	Thr	Asn	Ala	Phe	Ser	Thr	Lys	Asn	Val	Ala	Phe	Asn	His	1	5	10	15
Asp	Ser	Ser	Ser	Phe	Asp	His	Thr	Val	Asp	Gly	Phe	Leu	Thr	Ala	Asp	20	25	30	
Thr	Trp	Tyr	Arg	Pro	Lys	Ser	Ile	Leu	Ala	Asn	Gly	Thr	Thr	Trp	Arg	35	40	45	
Asp	Ser	Thr	Asp	Lys	Asp	Met	Arg	Pro	Leu	Ile	Thr	Val	Trp	Trp	Pro	50	55	60	
Asn	Lys	Asn	Val	Gln	Val	Asn	Tyr	Leu	Asn	Phe	Met	Lys	Ala	Asn	Gly	65	70	75	80
Leu	Leu	Thr	Thr	Ala	Ala	Gln	Tyr	Thr	Leu	His	Ser	Asp	Gln	Tyr	Asp	85	90	95	
Leu	Asn	Gln	Ala	Ala	Gln	Asp	Val	Gln	Val	Ala	Ile	Glu	Arg	Arg	Ile	100	105	110	
Ala	Ser	Glu	His	Gly	Thr	Asp	Trp	Leu	Gln	Lys	Leu	Leu	Phe	Glu	Ser	115	120	125	
Gln	Asn	Asn	Asn	Pro	Ser	Phe	Val	Lys	Gln	Gln	Phe	Ile	Trp	Asn	Lys	130	135	140	
Asp	Ser	Glu	Tyr	His	Gly	Gly	Gly	Asp	Ala	Trp	Phe	Gln	Gly	Gly	Tyr	145	150	155	160

WO 02/074943

PCT/FR02/00951

Ser Ile Pro Ala Asp Val Ala Ile Lys Gln Trp Ser Ala Lys Tyr Met
 2785 2790 2795 2800

Asn Gly Thr Asn Val Leu Gly Asn Gly Met Gly Tyr Val Leu Lys Asp
 2805 2810 2815

Trp His Asn Gly Gln Tyr Phe Lys Leu Asp Gly Asp Lys Ser Thr Leu
 2820 2825 2830

Pro Gln Ile
 2835

<210> 3

<211> 2568

<212> DNA

<213>

Nucleotides (catalytic domain n° 2)

<400> 3

```

gatatgtcta ctaatgcttt ttctacaaaa aatgttgcat tcaatcatga cagtagcagt 60
ttcgaccata ctgttgatgg ctctctgacg gcagatactt ggtatcgacc aaagtcaatt 120
ttggctaacg ggacaacttg gcgtgattcg actgataagg atatgcgacc attaatacact 180
gtttgggtggc caaataagaa tgttcaagtc aactacctca acttcatgaa agcaaattggc 240
ttgttgacaa cagcagcaca atacacacta cattcagatc aatatgattt gaaccaagct 300
gcacaagatg ttcaagtggc cattgaaagg cgcatcgctg cagagcatgg cacagactgg 360
ttacagaaat tgttgtttga atcacaaaat aataacccat catttgtgaa gcaacaattc 420
atttggaaaca aggattctga atatcatggt ggtgggtgatg ctgggttcca aggtggttat 480
ctgaagtatg gcaataaccc actcacacca acaactaatt ctgattatcg tcaacctggt 540
aatgcatttg atttcttgct agccaacgac gtggataatt ctaatcctgt tgtgcaagct 600
gaaaacttaa actggttaca ttacttaatg aactttggca ccatcactgc gggtaagat 660
gacgctaatt ttgatagtat tcgtattgac gctgctgact ttattcataa tgatacaatc 720
caacgtactt atgattatct tcgtgatgct tatcaagtgc aacaaagtga agccaaagca 780
aaccagcaca tttcattgggt tgaagctggc tttagcgcag gtacatcaac gattcataat 840
gatgcgttaa ttgagtcaaa cctccgtgaa gcagcgacat tgcgttaac aaatgaacct 900
ggtaaaaaata aaccattgac gaatatgcta caagacgttg acggcgggtac gcttatcacc 960
gaccatacgc agaatagtac agaaaatcag gcgacaccaa actattcaat tattcacgcg 1020
cacgataaag gtgtgcaaga aaaagtaggt gcagccatta ctgatgctac tgggtgctgat 1080
tggacgaact ttacagatga acagttaaaa gccggattag agctattcta taaggatcag 1140
cgcgcaacaa acaaaaagta taatagttat aacataccaa gtatttatgc cctgatgttg 1200
acaaacaaag atactgttcc tcgtatgtat tatggggata tgtatcaaga tgacggacag 1260
tatatggcaa acaagagtat ctactatgat gccttagtgt cattaatgac ggctcgtaaa 1320
agctatgtca gcggtggtca aactatgagt gttgacaatc atgggttggt gaagagtgtc 1380
cgttttggaa aagatgcgat gacagctaat gatttaggta catcagctac gcgtactgag 1440
ggctcttggtg tcattattgg taatgatcca aagttgcaac ttaatgattc ggataaagtg 1500
acactggata tgggtgcagc acataaaaaat caaaagtatc gcgcagttat cttacaaca 1560
cgtgatgggt tggcaacctt taattcagat caagcaccaa cagcttggac aaacgatcaa 1620
ggaacgttaa cattctcaaa tcaagagatt aacgggcaag acaatacaca aattcgtggt 1680
gttgctaate cgcaagtttc tggttatcta gctgtttggg tgcctgtggg tgcatacagac 1740
aatcaagatg cccgtacagc agcaacgaca acagaaaatc atgatggtaa agtattacac 1800
tcgaatgcgg cattagattc taaccttatt tatgaagggt tctctaactt ccaacctaa 1860
gcaacaacgc atgatgaact tacgaacgtt gtaattgcta aaaatgccga tgtcttcaat 1920
aattggggta ttacgagttt tgaaatggca ccacagtacc gttcaagtgg ggaccataca 1980
ttcttggtt caacgattga taatggttat gccttactg atcgctatga cttaggtttc 2040
atacaccaa caaagtatgg cactgatggt gatttgcgtg caacgattca agcgctacat 2100
catgctaata tgcagattt ggctgacgtt gttgataacc aggtctataa cttacctggt 2160
aaagaagttg ttctagcaac acgagcaggt gtttatggta atgacgacgc cacgggcttt 2220
ggaacgcaac tctatgtgac taactccgtt ggtggtggtc aataccaaga gaaatatgct 2280
ggacaatact tagaagctct gaaagcaaag tatccagacc tctttgaggg taaggcctat 2340
gattattggt ataagaacta tgcaaatgat gggtaaaatc cttactatac attgtcacac 2400
ggtgaccgtg aatctatccc agcagatggt gctattaagc aatggtcagc taagtatatg 2460

```

WO 02/074943

PCT/FR02/00951

aacggcacga acgttttggg caatggtatg ggttatgtat tgaaggattg gcataatggt 2520
 caatatttca agcttgatgg tgataaatca acattacctc aaatttaa 2568

<210> 4
 <211> 8506
 <212> DNA
 <213> DSR-E coding sequence

<400> 4
 atgagagaca tgagggtaat ttgtgaccgt aaaaaattgt acaaatcggg caaagtacta 60
 gtaacagccg gtatttttgc tttgatgatg tttggcgtca caactgctag tgtttagtga 120
 aatacgattg cagttgacac gaatcatagc cgtacttcag cacagattaa taagagtgcc 180
 gttgataagg ttaatgatga caagactact ttaggagcgg caaaagtagt ggtagtagcc 240
 acaacgccag cgacaacggg agcagataaa acagtaagtg caccgcgagc agataaggca 300
 gtagatacaa cgtcatcaac gacacctgca acggataagg cagtagatac aacgccaaacg 360
 acacctgcag cagataaggc agtagataca acgccaacga cacctgcagc agataaggca 420
 gtagatacaa cgccaacgac acctgcagca aataaagcag tagatacaac gccagcgacc 480
 gctgcaacag ataaggcggg agccacgcca gccacacctg cagcagataa gctagcaaat 540
 acgacgcctg caacggacaa ggcagtagcc acaacgccag cgacgcccgt agcaaatata 600
 gcagcagaca cagatagat tcatgatcaa ccattagata caaatgtgcc aactgataaa 660
 tcagcaaac tctgtctgac aacacaaaaa agtaccgata atcaacaagt taagtctaca 720
 gaaacatctc atcttcaaga aatcaacggg aaaacctatt ttcttgacga caatggtcaa 780
 gttaaaaaga acttcaccgc tattattgac ggttaaagttc tatactttga taaaacatcc 840
 ggcgaattga ccgcaaatgc accgcaagtt actaagggat tagtaaatat tgataatgca 900
 cataacgcgg ctcatgatct ccatgtatg cacagctgat atgtcgatgg ttacttaaca 960
 gctaacagtt ggtatcgctc taaggacatc ttaaaaaacg gaacgacctg gacaccaaca 1020
 acagcagaag attttcgacc attgtlcatg tcttggtggc cggataagaa tacgcaggta 1080
 gcttatctac aatataatga atcagttggt atgtacctg acgatgttaa agtatcaaat 1140
 gatgataata tgagcacatt gactgatgct gctatgactg ttcaaaagaa tatcgatatc 1200
 cgaattgggtg tatctggaaa aactgattgg ctcaagcaag atatgaacaa actgattgat 1260
 tcacaggcaa atttgaatat tgatagttaa tcaaaagggt atgatcattt acagggtggg 1320
 gcattgttat atgtgaatga tgacaaaaca cctaacgcga actcagatta ccgtctgtta 1380
 aaccgtacac caaccaacca aaccggccaa attactgatc caagtaaaaca aggtggatat 1440
 gagatgttat tagctaata tggtgataat tctaaccctg ttgtacaagc tgagcaattg 1500
 aactggcttc actacatgat gaacattggt actatagctc agaacgaccc aacagctaata 1560
 tttgacgggt atcgtgttga tgcggtgatg aacgttgatg ccgatctctt acaaattgct 1620
 ggtgattact ttaaagctgc atacggtact ggtaaaactg aggcaaacgc aaacaatcat 1680
 atttcgatct tgggaagatt ggataataat gattctgcgt acattaaagc ccacgggaat 1740
 aaccaattga caatggattt tccagcacac ttggctttga aatacgcctt gaacatgcct 1800
 cttgccgcac aaagtggcct agaaccgcta attaatatac gtcttggtta gcgtgggaaa 1860
 gatgccacag aaaatgaagc acaaccaaac tatgccttta tccgtgcccc tgatagttaa 1920
 gtgcagaccg ttattgcaca aattattaag gataaaatta acacaaaatc agacggctta 1980
 actgtaacac cagatgagat taagcaagct ttcactattt acaacgccga tgaattaaaa 2040
 gcagataagg aatatacagc atacaatatt cctgcttctt acgctgtatt gttgacaaac 2100
 aaggatactg tgccacgtgt ttattatggt gatctatttt ctgatgatgg acagtatatg 2160
 tcacagaagt caccatacta tgacgccatt acgtcacttt tgaaaagccg tatcaaatat 2220
 gttgctgggt gtcaaagtat gaatatgacg tacttgcatg agtgctttga tccagcaaaa 2280
 aatgagacaa agccacaagg tgtcttaaca tcagtacgtt acggtaaagg tgcgatgacg 2340
 gctgacgatt tgggtaatat tgacacacgt caacaaggta ttggtttggt gattaataat 2400
 aagccattct tgaattttaa tgatgatgaa caaattgtgc tcaatatggg tgctgctcac 2460
 aaaaatcaag cttaccgacc acttatgttg acaacaaaat ctggtcttca aattttacgat 2520
 aaggatgccg gagcgccagt tgtttatact aacgatgctg gtcaacttat ttttaagtca 2580
 gatatggtct atggtgtcag caatccacag gtatctggtt attttgctgc atgggtacca 2640
 gtcggtgcga gtgatagtca agatgctaga acacaaagca gccagtcaga aactaaggat 2700
 ggcgatgtct atcaatcaaa tgctgcgctt gattctaatg tgatttatga aggcttctcg 2760
 aatttccaag caatgcctga aaagaatgat gacttcacca acgtaaaaat tgctcaaaaat 2820
 gctaaattgt ttaaagattt agggattaca agctttgaat tagcaccgca atatcgttca 2880
 agtacagata atagtttttt ggattcgggt atccaaaacg gctatgcctt tactgatcga 2940
 tatgatgttg gctataatac gccacaaaaa tatggtacag ttgatcaact tctagatagt 3000

WO 02/074943

PCT/FR02/00951

```

caacgtactt atgattatct tcgtgatgct tatcaagtgc aacaaagtga agccaaagca 6720
aaccagcaca ttccattggg tgaagctggc ttagacgcag gtacatcaac gattcataat 6780
gatgcgttaa ttgagtcaaa cctccgtgaa gcagcgacat tgctgtaaac aaatgaacct 6840
ggtaaaaaata aaccattgac gaatatgcta caagacgttg acggcggtac gcttatcacc 6900
gaccatacgc agaatagtac agaaaatcag gcgacaccaa actattcaat tattcacgcg 6960
cacgataaag gtgtgcaaga aaaaagtaggt gcagccatta ctgatgctac tgggtgctgat 7020
tggacgaact ttacagatga acagttaaaa gccggattag agctattcta taaggatcag 7080
cgcgcaacaa acaaaaagta taatagttat aacataccaa gtattttatgc cctgatgttg 7140
acaaacaaaag atactgttcc tcgtatgtat tatggggata tgtatcaaga tgacggacag 7200
tatattggcaa acaagagtat ctactatgat gccttagtgt cattaatgac ggctcgtaaa 7260
agctatgtca gcggtggtca aactatgagt gttgacaatc atggtttgtt gaagagtgtc 7320
cgttttggaa aagatgcgat gacagctaata gatttaggta catcagctac gcgtactgag 7380
ggctcttggtg tcattattgg taatgatcca aagttgcaac ttaatgattc ggataaagtg 7440
acactggata tgggtgcagc acataaaaaat caaaagtatc gcgcagttat cttacaaca 7500
cgtgatgggtt tggcaacctt taattcagat caagcaccaa cagcttggac aaacgatcaa 7560
ggaaacgttaa cattctcaaa tcaagagatt aacgggcaag acaatacaca aattcgtggt 7620
gttgctaate cgcaagtttc tggttatcta gctgtttggg tgcctgtggg tgcacagac 7680
aatcaagatg cccgtacagc agcaacgaca acagaaaatc atgatggtaa agtattacac 7740
tcgaatgcgg cattagattc taaccttatt tatgaaggtt tctctaactt ccaacctaa 7800
gcaacaacgc atgatgaact tacgaacgtt ataatgcta aaaatgccga tgtcttcaat 7860
aattggggta ttacgagttt tgaatggca ccacagtagc gttcaagtgg ggaccataca 7920
ttcttggtatt caacgattga taatgtttat gccttcaact atcgctatga cttagggttc 7980
aatacaccaa caaagtatgg cactgatggt gatttgcgtg caacgattca agcgctacat 8040
catgctaata tgcaagttat ggctgacgtt gttgataacc aggtctataa cttacctggt 8100
aaagaagttg tttcagcaac acgagcaggt gtttatggta atgacgacgc cacgggcttt 8160
ggaacgcaac tctatgtgac taactccgtt ggtggtggtc aataccaaga gaaatagtct 8220
ggacaatact tagaagctct gaaagcaaaag tatccagacc tctttgaggg taaggcctat 8280
gattattggt ataagaacta tgcaaatgat ggttcaaate cttactatac attgtcacac 8340
ggtgaccgtg aatctatccc agcagatggt gctattaagc aatggtcagc taagtatatg 8400
aacggcacga acgttttggg caatggtatg ggttatgtat tgaaggattg gcataatggt 8460
caatatttca agcttgatgg tgataaatca acattacctc aaattt 8506

```

<210> 5
<211> 8931
<212> DNA
<213> Dsr-E gene

```

<400> 5
aataatctgt ctccattgct ttcaaaataa taatagttaa ttattatcat ggaacaatca 60
atattttatt tatattcaact attgaatata cttttttgca taaatctcta gagccgattt 120
tttggtttat acaatgaatt ggtaaagggt aatcattttt acaaaacat ggtggttttt 180
tattttttct aaaattaccg aactagagga agagaaaagg agcaatagtt gtatgagaga 240
catgagggta atttgtgacc gtaaaaaatt gtacaaatcg ggcaaagtac tagtaacagc 300
cggatttttt gctttgatga tgtttggcgt cacaactgct agtgttagtg caaatacgt 360
tgcagttgac acgaatcata gccgtacttc agcacagatt aataagagtg ccgttgataa 420
ggttaatgat gacaagacta ctttaggagc ggcaaaagta gtggcagtag ccacaacgcc 480
agcgacaccg gtagcagata aaacagtaag tgcaccgca gcagataagg cagtagatac 540
aacgtcatca acgacacctg caacggataa ggcagtagat acaacgccaa cgacacctgc 600
agcagataag gcagtagata caacgccaac gacacctgca gcagataagg cagtagatac 660
aacgccaacg acacctgcag caaataaagc agtagataca acgccagcga ccgctgcaac 720
agataaggcg gtagccacgc cagccacacc tgcagcagat aagctagcaa atacgacgcc 780
tgcaacggac aaggcagtag ccacaacgcc agcgacgccg gtagcaaata aagcagcaga 840
cacgagtagt attcatgatc aaccattaga tacaatgtg ccaactgata aatcagcaaa 900
cctcgctcgc acaacacaaa aaagtacgga taatcaacaa gttaagtcta cagaacatc 960
tcatcttcaa gaaatcaacg gtaaaaccta ttttcttgac gacaatggc aagttaaaaa 1020
gaacttcacc gctattattg acggtaaagt tctatacttt gataaaacat ccggcgaatt 1080
gaccgcaaat gcaccgcaag ttactaaggg attagtaaat attgataatg cacataacgc 1140
ggctcatgat ctcacagctg ataactcac aaatgtcgat gggtacttaa cagctaacag 1200
ttggtatcgt cctaaggaca tcttaaaaaa cggaacgacc tggacaccaa caacagcaga 1260

```